Prof. Philip Koopman

Safety Performance Indicators (SPIs) for Self-Driving Cars

June 2020

Carnegie Mellon University

@PhilKoopman

EDGE CASE RESEARCH

Overview

- KPIs: Key Performance Indicators
  - Quantify performance
  - Important, but not enough for safety

- SPIs: Safety Performance Indicators
  - Quantify safety
  - Leading vs. Lagging SPIs
  - Safety case validity SPIs

https://on.gei.co/2r2rjzg

Key Performance Indicator (KPI)

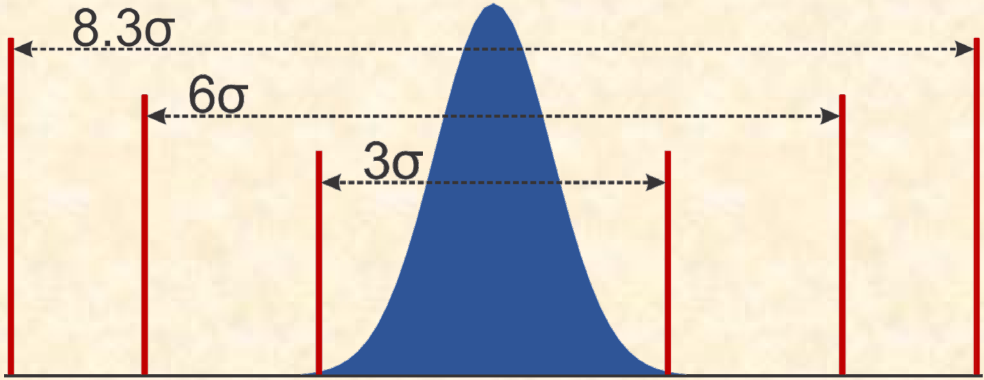
EDGE CASE RESEARCH

- KPI:
  - Quantifiable measurement
  - Used to gauge statistical performance
- KPI examples:
  - Percent correctly identified pedestrians
  - Miles between SDC self-disengagements
  - Miles between uncomfortable braking
- KPIs can measure SDC progress
  - Metrics should improve over time
  - But – KPIs are wrong approach for safety

https://bit.ly/2ZQcIYC

Six Sigma Isn't Enough for Safety

- KPIs help with quality
  - Are all functions working?
  - Is the functionality improving?
  - Is the fault rate decreasing?



- Good KPIs are only the start
  - Six Sigma Quality: 99.99966% (five nines)
    - A good start; not enough for life critical functions
  - Fatal Crash Avoidance: 99.99999999996% (eleven nines)
    - Safety is 1 million times more demanding! ➔ 8.34 sigma
      - » (example: 1000 opportunities/mile, 250M miles/fatal crash, 1.5σ shift)

Functionality vs. Safety

- Functionality (KPIs):
  - Are all the features implemented?
  - Does each feature work as intended?
  - Are all scenarios accounted for?
  - Does the product do what it is supposed to?

https://bit.ly/2MaLkfY

- Safety:
  - Are there dangerous mis-behaviors?
  - Are there dangerous gaps in the Operational Design Domain?
  - Are there dangerous gaps in fault responses?
  - Are there dangerous defects in requirements, design, repair, etc.?

Safety Performance Indicator (SPI)

- SPI:
  - Quantifiable measurement
  - Used to gauge safety
  - Typically:
    arrival rate of adverse events
    compared to a risk budget

https://bit.ly/2MaLkfY

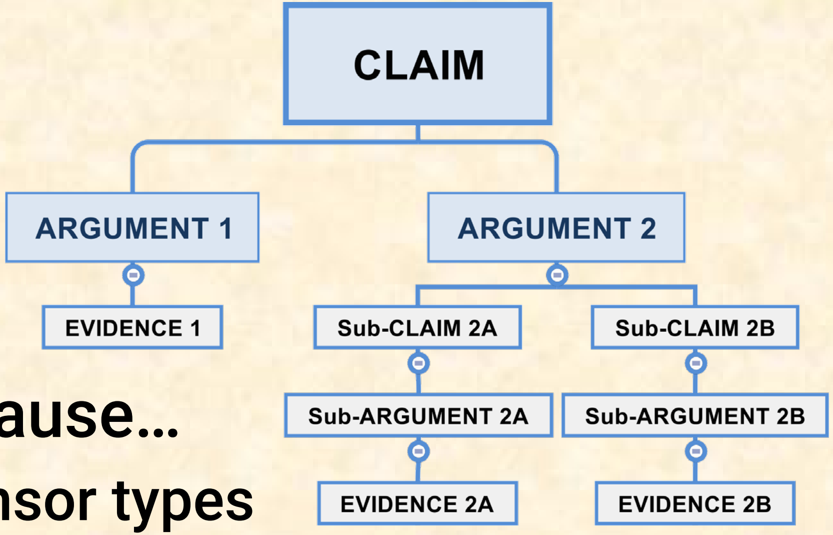
- Lagging SPI metrics:
  (per hour is implied)
  - Loss events (crashes) per hour
  - Incidents (could have been a loss event)
    - Example: running a red light, driving wrong direction for lane

Leading SPIs

- System Level Leading SPIs:
  - Road test incidents caught by safety driver
  - Simulator (SIL/HIL) incidents
- Subsystem Leading SPIs:
  - Vehicle Controls: compromised vehicle stability
  - Path Planning: insufficient clearance to object
  - Perception: false negative (non-detection)
  - Prediction: unexpected object behavior
- Lifecycle SPIs:
  - Maintenance errors
  - Invalid configuration installed

Safety Case

- System is safe because …
  - Explanation of why
  - Evidence supporting explanation
  - Assumptions



- Ex.: SDC misses pedestrians because…
  - Pedestrians are detected with 3 sensor types
  - Pedestrian intent is predicted accurately
  - Path planning leaves buffer zone around them

- SPIs help detect violations of the safety case

SPIs and the Safety Case

- SPIs also measure safety case assumptions
  - ODD matches the Operational Domain
  - Validation predicts operational performance
  - Maintenance performed as required
  - Correct configuration installed in vehicle

- Example Safety Case-related SPIs:
  - Appearance of assumed rare objects and events
  - Correlated diverse sensor detection faults
  - Safety related maintenance error

https://bit.ly/3gHWiYu

KPI vs. SPI Contrast

Distance to object:

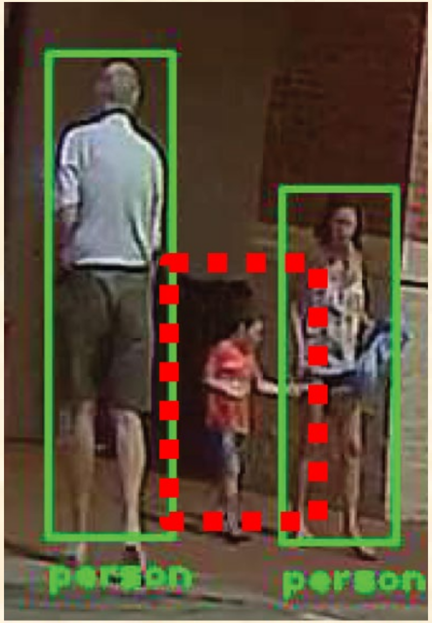
- KPI: average and 95th percentile clearance
- SPI: how often SDC violates safe clearance limit

Sensor effectiveness:

- KPI: detection rate, SNR per sensor
- SPI: concurrent multi-sensor detection failure
- SPI: loss of calibration

Pedestrian perception:

- KPI: accuracy, precision, recall
- SPI: false negative for more than <k> consecutive frames
- SPI: previously unknown type of pedestrian encountered



SPIs and the Deployment Decision

EDGE CASE RESEARCH

- KPIs can predict if your SDC will “work”
  - SOTIF analysis resolves many outliers
- SPIs can predict if it will work safely
  - System level SPIs from simulation & testing
    - At system level, an outlier could be fatal
  - Subsystem SPIs
    - Control, planning, prediction, perception performance SPIs
    - Ability of system to detect and respond to exiting ODD
  - Safety case SPIs
    - Arrival rate of “surprises” / unknown unknowns during testing
    - Arrival rate of gaps in safety case being discovered

Conclusions

- SPIs predict and monitor system safety
  - KPIs: “how well do we drive”
  - SPIs: “how often are we potentially unsafe”
- Different flavors of SPIs
  - Lagging (e.g., crash rates)
  - Leading (e.g., simulator collisions, testing incidents)
  - Safety case SPIs (how often is safety case invalid)
- Do you have SPI coverage for your system?
  - Extend SOTIF analysis beyond KPIs to include SPIs
  - See ANSI/UL 4600 Chapter 16 on SPIs

EDGE CASE RESEARCH

WE DELIVER THE PROMISE OF AUTONOMY